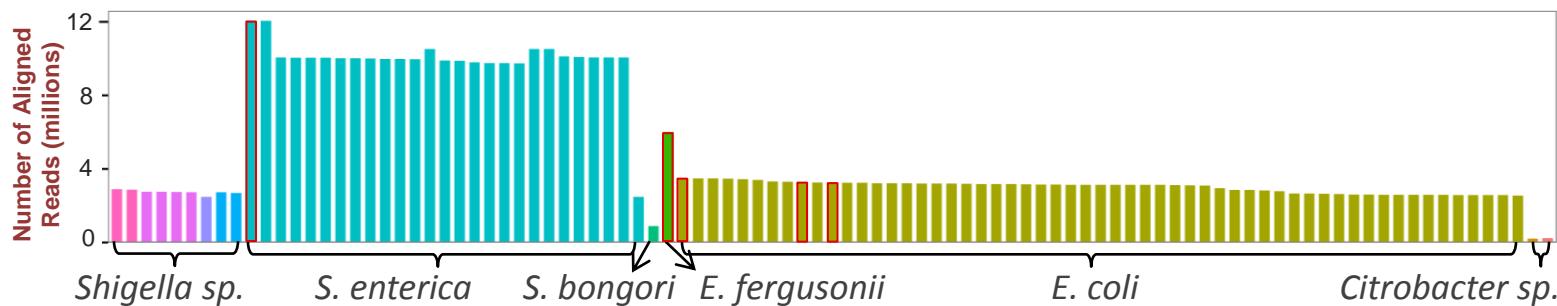
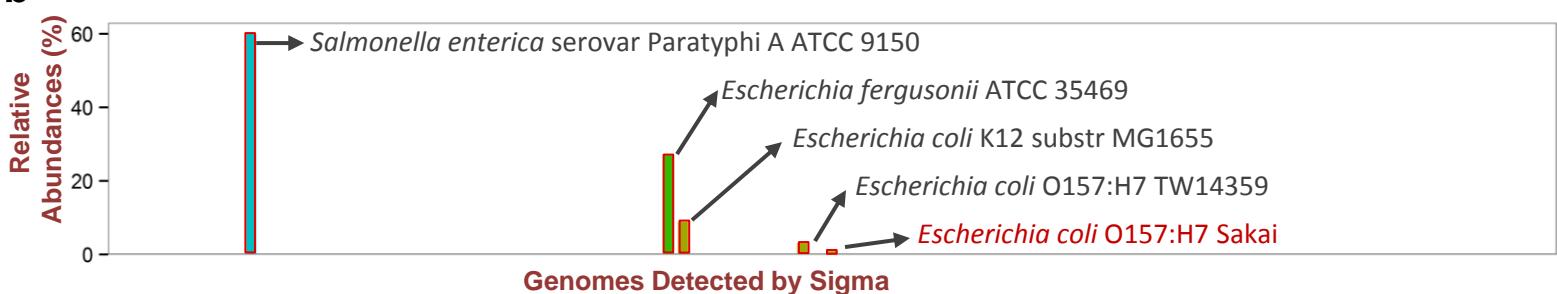


Supplementary Figure 1. Detection of an *E. coli* strain from a 5-genome synthetic community. (a) Alignment of reads to many genomes related to the five mixed genomes highlighted in red outline. (b) Accurate estimation of the relative abundances of the five correctly identified genomes by Sigma. The genomes are listed in the same order in the two bar charts.

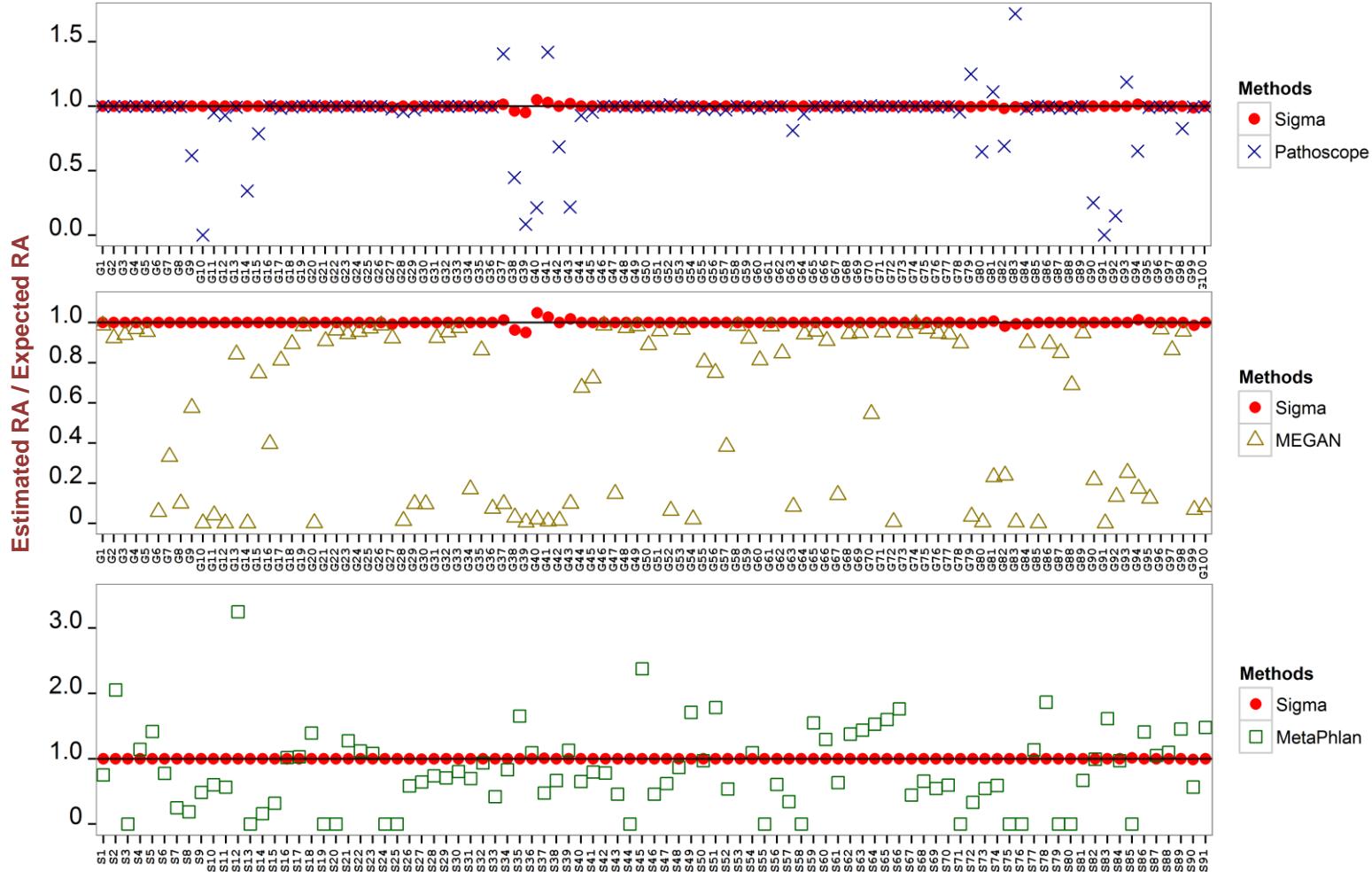
a



b



Supplementary Figure 2. Comparison of Sigma with Pathoscope, MetaPhlAn and MEGAN on a 100-genome synthetic community. The quantification accuracy of an algorithm is measured by the ratios of the estimated relative abundances (RA) to the expected RA (i.e. a ratio of 1 for perfect estimation) across the 100 genomes. The genomes are listed alphabetically in the x-axis labeled by their genome or species indices. The legend shows the names and the expected RA of the 100 genomes. Species represented by multiple genomes are highlighted in red in the legend.



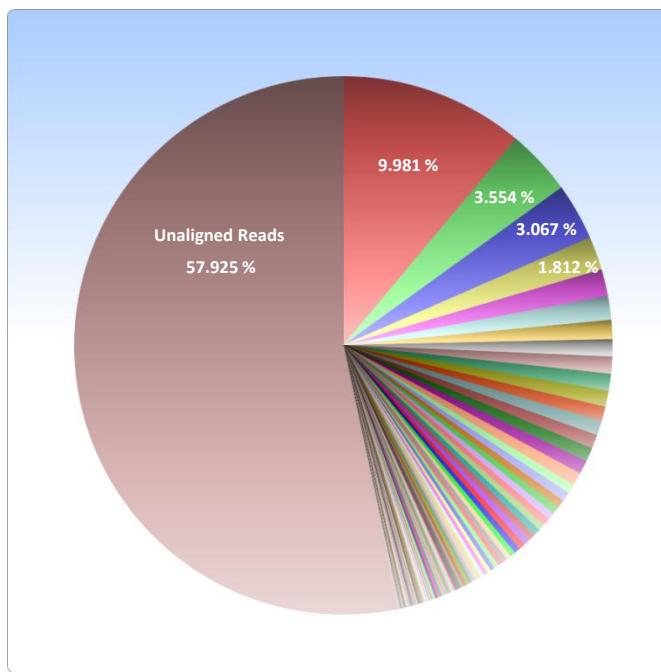
Genome Index (Species Index)	Geome Name	Expected RA	Genome Index (Species Index)	Geome Name	Expected RA	Genome Index (Species Index)	Geome Name	Expected RA
G1(S1)	<i>Acidobacterium capsulatum</i> ATCC 51196	0.52	G35(S35)	<i>Dictyoglomus turgidum</i> DSM 6724	0.51	G69(S62)	<i>Pyrobaculum aerophilum</i> IM2	0.99
G2(S2)	<i>Aciduliprofundum boonei</i> T469	0.59	G36(S36)	<i>Enterococcus faecalis</i> OG1RF	0.99	G70(S63)	<i>Pyrobaculum arsenaticum</i> DSM 13514	0.84
G3(S3)	<i>Aerococcus urinae</i> ACS 120 V Col10a	0.75	G37(S37)	<i>Escherichia coli</i> 042	0.77	G71(S64)	<i>Pyrobaculum calidifontis</i> JCM 11548	0.7
G4(S4)	<i>Akkermansia muciniphila</i> ATCC BAA 835	0.51	G38(S37)	<i>Escherichia coli</i> ATCC 8739	0.54	G72(S65)	<i>Pyrococcus furiosus</i> DSM 3638	0.61
G5(S5)	<i>Archaeoglobus fulgidus</i> DSM 4304	0.96	G39(S37)	<i>Escherichia coli</i> K 12 substr MG1655	0.61	G73(S66)	<i>Pyrococcus horikoshii</i> OT3	0.46
G6(S6)	<i>Bacillus amyloliquefaciens</i> DSM 7	1.33	G40(S37)	<i>Escherichia coli</i> O104 H4 2011C 3493	0.41	G74(S67)	<i>Rhodopirellula baltica</i> SH 1	0.82
G7(S7)	<i>Bacillus cereus</i> ATCC 10987	1.3	G41(S37)	<i>Escherichia coli</i> O55 H7 C89615	1.09	G75(S68)	<i>Ruegeria pomeroyi</i> DSS 3	1.28
G8(S8)	<i>Bacillus thuringiensis</i> BMB171	1.41	G42(S37)	<i>Escherichia coli</i> O7 K1 CE10	0.66	G76(S69)	<i>Salinisporea arenicola</i> CNS 205	0.63
G9(S9)	<i>Bacteroides thetaiotaomicron</i> VPI 5482	1.56	G43(S37)	<i>Escherichia coli</i> UMN026	0.56	G77(S70)	<i>Salinisporea tropica</i> CNB 440	1.36
G10(S10)	<i>Bacteroides vulgaris</i> ATCC 8482	1.05	G44(S38)	<i>Escherichia fergusonii</i> ATCC 35469	0.86	G78(S71)	<i>Salmonella bongori</i> NCTC 12419	0.48
G11(S11)	<i>Bordetella bronchiseptica</i> RB50	0.52	G45(S39)	<i>Fusobacterium nucleatum</i> ATCC 25586	1.16	G79(S72)	<i>Salmonella enterica</i> serovar Paratyphi A ATCC 9150	1.56
G12(S12)	<i>Buchnera aphidicola</i> TLW03 <i>Acyrhosiphon pisum</i>	0.75	G46(S40)	<i>Gemmatumonas aurantiaca</i> T 27	0.5	G80(S72)	<i>Salmonella enterica</i> serovar Typhimurium	1.04
G13(S13)	<i>Burkholderia</i> 383	0.52	G47(S41)	<i>Geobacter sulfurreducens</i> PCA	0.48	G81(S73)	<i>Shewanella baltica</i> OS185	1.55
G14(S14)	<i>Burkholderia pseudomallei</i> 668	1.25	G48(S42)	<i>Haloferax volcanii</i> DS2	1.07	G82(S73)	<i>Shewanella baltica</i> OS223	0.56
G15(S15)	<i>Burkholderia xenovorans</i> LB400	1.48	G49(S43)	<i>Herpetosiphon aurantiacus</i> DSM 785	0.84	G83(S74)	<i>Shigella flexneri</i> 2002017	1.59
G16(S16)	<i>Caldicellulosiruptor bescii</i> DSM 6725	1.44	G50(S44)	<i>Hydrogenobaculum</i> YO4AAS1	1.14	G84(S75)	<i>Streptomyces venezuelae</i> ATCC 10712	1.58
G17(S17)	<i>Caldicellulosiruptor saccharolyticus</i> DSM 8903	1.33	G51(S45)	<i>Ignicoccus hospitalis</i> KIN4 I	1.32	G85(S76)	<i>Sulfobacillus acidophilus</i> DSM 10332	1.18
G18(S18)	<i>Calditerrivibrio nitroreducens</i> DSM 19672	0.71	G52(S46)	<i>Klebsiella pneumoniae</i> 1084	1.16	G86(S77)	<i>Sulfobolbus tokodaii</i> 7	0.98
G19(S19)	<i>Candidatus Accumulibacter phosphatis</i> clade IIA UW 1	1.4	G53(S47)	<i>Leptothrix choliniae</i> SP 6	0.49	G87(S78)	<i>Sulfurihydrogenibium azorense</i> Az Fu1	1.51
G20(S20)	<i>Candidatus Tremblaya princeps</i> PC1	1.13	G54(S48)	<i>Listeria monocytogenes</i> ATCC 19117	1.16	G88(S79)	<i>Sulfurihydrogenibium</i> YO3AOP1	0.69
G21(S21)	<i>Carboxydothermus hydrogenoformans</i> Z 2901	0.86	G55(S49)	<i>Methanocaldooccus jannaschii</i> DSM 2661	1.59	G89(S80)	<i>Thermincola potens</i> JR	1.42
G22(S22)	<i>Chlorobium limicola</i> DSM 245	1.01	G56(S50)	<i>Methanococcus maripaludis</i> C5	1.58	G90(S81)	<i>Thermoanaerobacter pseudethanolicus</i> ATCC 33223	1.21
G23(S23)	<i>Chlorobium phaeobacteroides</i> DSM 266	0.48	G57(S50)	<i>Methanococcus maripaludis</i> S2	1	G91(S82)	<i>Thermobaculum terrenum</i> ATCC BAA 798	1.46
G24(S24)	<i>Chlorobium phaeovibrioides</i> DSM 265	0.42	G58(S51)	<i>Methanopyrus kandleri</i> AV19	0.5	G92(S83)	<i>Thermotoga neapolitana</i> DSM 4359	1.16
G25(S25)	<i>Chlorobium tepidum</i> TLS	1.33	G59(S52)	<i>Methanoscincus acetivorans</i> C2A	1.11	G93(S84)	<i>Thermotoga petrophila</i> Rku 1	1.46
G26(S26)	<i>Chloroflexus aurantiacus</i> J 10 fl	0.54	G60(S53)	<i>Nanoarchaeum equitans</i> Kin4 M	1.51	G94(S85)	<i>Thermotoga</i> RQ2	0.74
G27(S27)	<i>Citrobacter koseri</i> ATCC BAA 895	1.31	G61(S54)	<i>Nitrosomonas europaea</i> ATCC 19718	1.49	G95(S86)	<i>Thermus thermophilus</i> HB8	1.26
G28(S28)	<i>Clostridium botulinum</i> A ATCC 19397	0.82	G62(S55)	<i>Nostoc</i> PCC 7120	0.88	G96(S87)	<i>Thiobacillus denitrificans</i> ATCC 25259	1.39
G29(S29)	<i>Clostridium difficile</i> 630	1.21	G63(S56)	<i>Pantoea ananatis</i>	1.54	G97(S88)	<i>Treponema denticola</i> ATCC 35405	1.34
G30(S30)	<i>Clostridium thermocellum</i> ATCC 27405	1.21	G64(S57)	<i>Pectobacterium carotovorum</i> PC1	1.19	G98(S89)	<i>Wolinella succinogenes</i> DSM 1740	1.22
G31(S31)	<i>Cytophaga hutchinsonii</i> ATCC 33406	0.52	G65(S58)	<i>Pelodictyon phaeoclastratiforme</i> BU 1	0.78	G99(S90)	<i>Yersinia pestis</i> CO92	1.51
G32(S32)	<i>Deinococcus radiodurans</i> R1	1.02	G66(S59)	<i>Persephonella marina</i> EX H1	0.95	G100(S91)	<i>Zymomonas mobilis</i> ZM4	0.89
G33(S33)	<i>Desulfovibrio desulfuricans</i> ATCC 27774	0.53	G67(S60)	<i>Porphyromonas gingivalis</i> ATCC 33277	1.19			
G34(S34)	<i>Desulfovibrio vulgaris</i> DP4	0.64	G68(S61)	<i>Providencia stuartii</i> MRSN 2154	0.49			

Supplementary Figure 3. Taxonomic composition of a fecal metagenome estimated by Sigma using 24,994 reference genomes from GenBank. (a) Names and relative abundances of identified genomes. (b) Pie chart of relative abundances of identified genomes and unaligned reads. (c) Pie chart of percentages of aligned reads assigned to identified genomes.

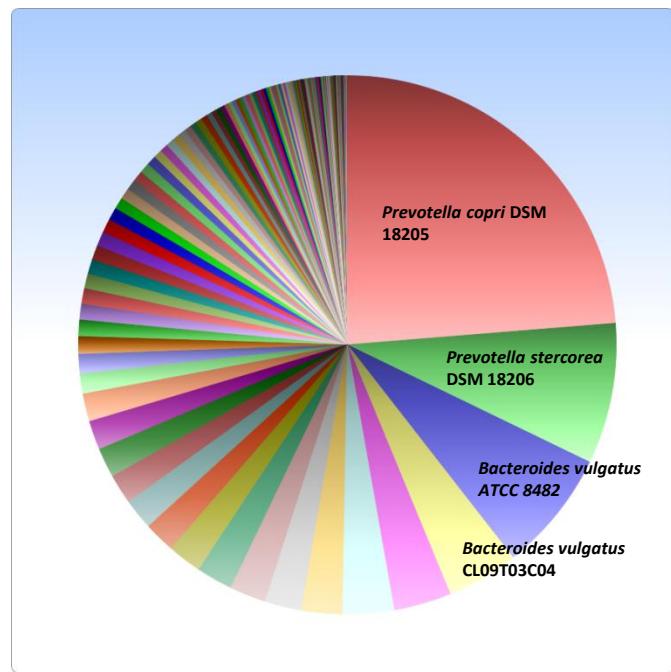
a

Genome Name	RA (%)	Genome Name	RA (%)	Genome Name	RA (%)	Genome Name	RA (%)
<i>Prevotella copri</i> DSM 18205	9.981	<i>Parabacteroides</i> sp. 20_3	0.244	<i>Parabacteroides</i> sp. D13	0.059	<i>Clostridium leptum</i> DSM 753	0.019
<i>Prevotella stercorea</i> DSM 18206	3.554	<i>Bacteroides</i> sp. 3_1_23	0.215	<i>Roseburia intestinalis</i> L1-82	0.055	<i>Bacteroides salyersiae</i> WAL 10018	0.019
<i>Bacteroides vulgatus</i> ATCC 8482	3.067	<i>Bacteroides dorei</i> CL02T12C06	0.208	<i>Bacteroides xylanisolvans</i> SD CC 1b	0.054	<i>Clostridium</i> sp. JCC	0.018
<i>Bacteroides vulgatus</i> CL09T03C04	1.812	<i>Bacteroides coprophilus</i> DSM 18228	0.198	<i>Alistipes finegoldii</i> DSM 17242	0.048	<i>Blautia wexlerae</i> DSM 19850	0.018
<i>Bacteroides vulgatus</i> dnLV7	1.454	<i>Bacteroidales bacterium</i> ph8	0.196	<i>Roseburia intestinalis</i> XB684	0.045	<i>butyrate-producing bacterium</i> SS3/4	0.018
<i>Bacteroides xylanisolvans</i> XB1A	1.285	<i>Clostridium</i> sp. L2-50	0.180	<i>Bacteroides</i> sp. 2_1_33B	0.044	<i>Prevotella bergensis</i> DSM 17361	0.018
<i>Bacteroides</i> sp. 3_1_40A	1.031	<i>Butyrimonas virosa</i> DSM 23226	0.179	<i>Bacteroides fragilis</i> str. Korea 419	0.043	<i>Parabacteroides</i> sp. D25	0.018
<i>Bacteroides</i> sp. 4_3_47FAA	0.929	<i>Capnocytophaga</i> sp. oral taxon 329 str. F0087	0.166	<i>Roseburia intestinalis</i> M50/1	0.041	<i>Bifidobacterium adolescentis</i> L2-32	0.017
<i>Bacteroides ovatus</i> SD CMC 3f	0.922	<i>Odoribacter splanchnicus</i> DSM 220712	0.163	<i>Faecalibacter</i> cf. <i>prausnitzii</i> KLE1255	0.039	<i>Bacteroides</i> sp. D2	0.017
<i>Bacteroides uniformis</i> ATCC 8492	0.918	<i>Bacteroides xylanisolvans</i> CL03T12C04	0.158	<i>Alistipes senegalensis</i> JC50	0.038	<i>Dorea formicigenerans</i> 4_5_53AFAA	0.017
<i>Bacteroides</i> sp. 1_1_30	0.862	<i>Bacteroides</i> sp. 9_1_42FAA	0.137	<i>Bacteroides massiliensis</i> dnLV3	0.038	<i>Parabacteroides distasonis</i> CL03T12C09	0.016
<i>Bacteroides uniformis</i> CL03T12C37	0.790	<i>Bacteroides xylanisolvans</i> SD CC 2a	0.125	<i>Alistipes onderdonkii</i> DSM 19147	0.037	<i>Parabacteroides distasonis</i> ATCC 8503	0.016
<i>Bacteroides vulgatus</i> PCS10	0.782	<i>Eubacterium eligens</i> ATCC 27750	0.121	<i>Faecalibacterium prausnitzii</i> M21/2	0.034	<i>Catenibacterium mitsuokai</i> DSM 15897	0.016
<i>Parabacteroides merdae</i> ATCC 43184	0.753	<i>Bacteroides massiliensis</i> 884634	0.119	<i>Bacteroides cellulosilyticus</i> CL02T12C19	0.033	<i>Ruminococcus</i> sp. 5_1_39BFAA	0.015
<i>Parabacteroides merdae</i> CL09T00C40	0.749	<i>Bacteroides</i> sp. 3_1_33FAA	0.106	<i>Bacteroides fragilis</i> str. 3397 T10	0.033	<i>Barnesiella intestinihominis</i> YIT 11860	0.015
<i>Bacteroides</i> sp. 1_1_6	0.735	<i>Eubacterium ventriosum</i> ATCC 27560	0.102	<i>Bacteroides ovatus</i> 3_8_47FAA	0.033	<i>Clostridium</i> sp. M62/1	0.015
<i>Dialister invisus</i> DSM 15470	0.701	<i>Bacteroides pectinophilus</i> ATCC 43243	0.098	<i>Bacteroides massiliensis</i> B84634	0.032	<i>Bacteroides cellulosilyticus</i> DSM 14838	0.014
<i>Bacteroides thetaiotomicron</i> VPI-5482	0.504	<i>Bacteroides fragilis</i> 638R	0.093	<i>Bacteroides</i> sp. 3_1_19	0.032	<i>Prevotella multisaccharivorax</i> DSM 17128	0.014
<i>Bacteroides stercoris</i> ATCC 43183	0.492	<i>Eubacterium rectale</i> DSM 17629	0.091	<i>Ruminococcaceae bacterium</i> D16	0.032	<i>Porphyromonas bennonii</i> JCM 16335	0.013
<i>Parabacteroides merdae</i> CL03T12C32	0.444	<i>Paraprevotella xylaniphila</i> YIT 11841	0.090	<i>Bacteroides caccae</i> CL03T12C61	0.031	<i>Parabacteroides johnsonii</i> DSM 18315	0.013
<i>Bacteroides</i> sp. D20	0.415	<i>Bacteroides dorei</i> DSM 17855	0.090	<i>Faecalibacterium prausnitzii</i> SL3/3	0.031	<i>Bacteroides cellulosilyticus</i> WH2	0.013
<i>Bacteroides</i> sp. D22	0.397	<i>Roseburia inulinivorans</i> DSM 16841	0.081	<i>Prevotella</i> sp. BV3P1	0.029	<i>Bacteroides</i> sp. 14(A)	0.013
<i>Bacteroides faecis</i> MAJ27	0.395	<i>Eubacterium rectale</i> ATCC 33656	0.079	<i>Ruminococcus callidus</i> ATCC 27760	0.029	<i>Sutterella wadsworthensis</i> 3_1_45B	0.013
<i>Paraprevotella clara</i> YIT 11840	0.372	<i>Bacteroides dorei</i> CL03T12C01	0.073	<i>Faecalibacterium prausnitzii</i> A2-165	0.029	<i>Oscillibacter</i> sp. KLE 1745	0.012
<i>Bacteroides uniformis</i> dnLV2	0.369	<i>Bacteroides ovatus</i> CL02T12C04	0.072	<i>Ruminococcus torques</i> L2-14	0.028	<i>Bacteroides fragilis</i> str. DS-71	0.012
<i>Bacteroides</i> sp. 17565	0.366	<i>Eubacterium rectale</i> M104/1	0.071	<i>Faecalibacterium prausnitzii</i> L2-6	0.026	<i>Dorea</i> sp. AGR2135	0.012
<i>Alistipes putredinis</i> DSM 17216	0.364	<i>Roseburia hominis</i> A2-183	0.070	<i>Coprococcus comes</i> ATCC 27758	0.024	<i>Ruminococcus bromii</i> L2-63	0.012
<i>Bacteroides stercoris</i> CC31F	0.345	<i>Alistipes indistinctus</i> YIT 12060	0.068	<i>Subdoligranulum</i> sp. 4_3_54A2FAA	0.023	<i>Prevotella bivia</i> DSM 20514	0.011
<i>Bacteroides</i> sp. 1_1_14	0.323	<i>Alistipes shahii</i> WAL 8301	0.068	<i>Alistipes</i> sp. HGB5	0.023	<i>Prevotella timonensis</i> 4401737	0.011
<i>Bacteroides</i> sp. 4_1_36	0.287	<i>Bacteroides nordii</i> JCM 12987	0.064	<i>Bacteroides rodentium</i> JCM 16496	0.022	<i>Bacteroides fragilis</i> str. 3783N-8	0.011
<i>Bacteroides thetaiotomicron</i> dnLV9	0.277	<i>Bacteroides ovatus</i> CL03T12C18	0.062	<i>Bacteroides</i> sp. 2_1_16	0.021	<i>Collinsella aerofaciens</i> ATCC 25986	0.011
<i>Bacteroides</i> sp. D1	0.273	<i>Bacteroides</i> sp. CL09T03C10	0.062	<i>Dorea longicatena</i> DSM 13814	0.021	<i>Candidatus Alistipes marseilloanorexicus</i> AP11	0.010
<i>Bacteroides dorei</i> 5_1_36/D4	0.267	<i>Bacteroides ovatus</i> ATCC 8483	0.061	<i>Bacteroides uniformis</i> CL03T00C23	0.020	<i>Coprococcus catus</i> GD/7	0.010
<i>Bacteroides</i> sp. 2_2_4	0.259	<i>Dorea formicigenerans</i> ATCC 27755	0.059	<i>Sutterella wadsworthensis</i> HGA0223	0.019	<i>Bacteroides coprocola</i> DSM 17136	0.010

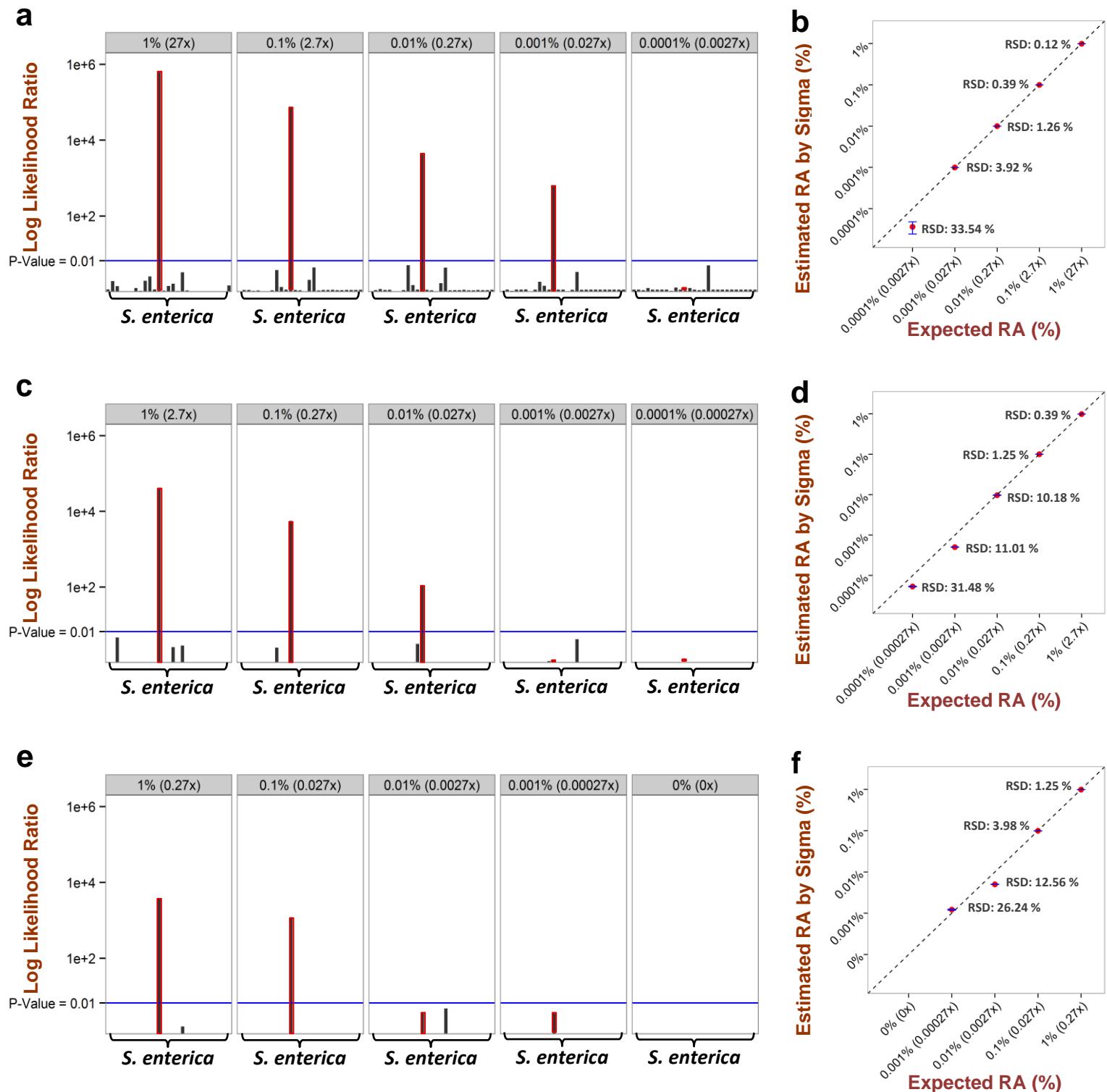
b



c



Supplementary Figure 4. Identification of a *Salmonella enterica* strain at a serial dilution of relative abundances in sub-samples of a human fecal metagenome. Various amounts of *Salmonella enterica* reads were spiked into the full dataset [(a) and (b)], a 10% sub-sample [(c) and (d)], and a 1% sub-sample [(e) and (f)] of a human fecal metagenome. (a), (c) and (e) show the likelihood ratios of all aligned *Salmonella enterica* strains in these samples. (b), (d), and (e) show the estimated and expected relative abundances (RA) of the spike-in *Salmonella enterica* strain in these samples.



Supplementary Table 1. Identification of a *S. enterica* strain in a human fecal microbiota background using different reference genome databases.

Test	Identified Genome	Relative Abundance (%)	[†] ANI % (SD %)	Genome Coverage (%)	High-confidence SNPs	[‡] Covered Region Divergence (%)
¹ Ref1	<i>Salmonella enterica</i> serovar Paratyphi A ATCC 9150	0.988		100	0	0
² Ref2	<i>Salmonella enterica</i> serovar Paratyphi A AKU 12601	0.986	99.99 (0.08)	99.95	190	0.004
³ Ref3	<i>Salmonella bongori</i> NCTC 12419	0.031	90.50 (3.39)	11.46	3750	0.734
⁴ Ref4	None					

[†] ANI % (SD %): average nucleotide identity and its standard deviation in percentage.

[‡] Coverage region divergence: the percentage of high-confidence SNPs out of the covered regions of the identified genome.

¹ Ref1 (2266 genomes): the full RefSeq database

² Ref2 (2265 genomes): the RefSeq database without the spike-in genome from *Salmonella enterica* serovar Paratyphi A ATCC 9150

³ Ref3 (2239 genomes): the RefSeq database without all *Salmonella enterica* species genomes

⁴ Ref4 (2238 genomes): the RefSeq database without all *Salmonella* genus genomes